

Khan, Paul (mayo 2006). *La información en la red hipertextual : El valor de las conexiones*. En: Encrucijadas, no. 37. Universidad de Buenos Aires. Disponible en el Repositorio Digital Institucional de la Universidad de Buenos Aires: <<http://repositorioubi.sisbi.uba.ar>>

La información en la red hipertextual

El valor de las conexiones *

El valor diferencial de la Web no está en encontrar cosas, sino en que los vínculos nos lleven a ellas, en una red hipertextual que podamos navegar. El entorno digital en el cual nos movemos hace que no pongamos atención en las conexiones más obvias. La trascendencia del problema no debe ser buscada en los pequeños sitios y páginas individuales, sino en los sitios de las grandes corporaciones, donde la ausencia de arquitecturas de la información cuidadosamente planificadas ocasiona que ciertos contenidos permanezcan escondidos en los espacios de información que albergan los sitios.

por Paul Khan

Especialista en Arquitectura de la Información, es profesor en el Master de Multimedia Hipermedia de la Universidad de París 1. Entre 1994 y 2001 dictó clases en la Escuela de Diseño de Rhode Island. Es coautor de los libros Mapping Web Sites y From Memex to Hypertext: Vannevar Bush and the Mind's Machine.

Máquinas imaginarias

El mayor valor que tiene la World Wide Web es proveernos la información que necesitamos, simplemente conectándonos a Internet. Esta información es conocida como "Web pública", conformada por millones de sitios, a los cuales podemos acceder desde nuestra oficina, hogar o comercio. La Web nos ofrece textos digitalizados, imágenes, sonidos, archivos de video conectados en una red hipertextual.

Una de las bases del concepto de hipertexto es el ensayo As We May Think [1], publicado en 1945, escrito por Vannevar Bush, profesor del MIT (Massachusetts Institute of Technology). Bush predijo las formas en que las máquinas podrían ayudar a los científicos a extender la capacidad de almacenar información y utilizarla cuando sus trabajos de investigación lo requirieran. Tres máquinas fueron descriptas en ese texto.

La primera era la "Vocoder", capaz de convertir la voz en texto, fue presentada en una demostración que tuvo lugar en Nueva York en 1939. El discurso es uno de los aspectos fundamentales de la cultura humana, pero transformar arbitrariamente la comunicación hablada en textos electrónicos coherentes y exactos continúa siendo un problema que no ha sido resuelto.

La segunda idea fue una cámara montada en la cabeza, graficada como un tercer ojo, pensada para permitirle al científico capturar imágenes de su trabajo en laboratorio sin necesidad de trasladarse o realizar demasiados movimientos. La predicción de Bush se materializa en la actualidad, entre otros ejemplos, en la inclusión de las cámaras digitales en los aparatos de telefonía celular.

La tercera idea de Bush inspiró a muchos teóricos e investigadores de las ciencias de la computación y nos proveyó de una de las primeras visiones de la Web que tenemos en la actualidad. Se trata de la "Memex", cuyo nombre evocaba al concepto de "Memoria extendida". La Memex estaba diseñada para almacenar una biblioteca completa de microfilms personalizada y estaba conformada por un escritorio con dos pantallas y un

teclado. Además, tenía una copiadora para agregar documentos, un dispositivo que permitía incluir imágenes en las pantallas.

Lo más trascendente en la “personalización” del sistema era que la lectora construía caminos (paths) de archivos asociados que podían ser almacenados y guardados en la máquina. Bush también imaginó que estos archivos podrían ser compartidos por diversos usuarios Memex.

En la actualidad, quien lea este ensayo debe comprender el contexto tecnológico de los años '30 y pasar por alto que muchos detalles de la Memex no fueron descriptos. El sueño de Bush era crear una red de almacenamiento e intercambio de archivos entre los intelectuales y científicos, que aseguraría que importantes investigaciones no se perdieran en los pasillos y estantes de las bibliotecas.

Conexiones perdidas

En ocasiones, resulta asombroso lo que se puede encontrar en la Web. Recientemente, estaba preparando una conferencia y decidí buscar información sobre Adolf Loos, influyente arquitecto y diseñador, al ver una placa de mármol ubicada afuera del edificio donde vivo.

Cuando vi la placa por primera vez, ya había escuchado sobre Tristan Tzara, quien creó el movimiento dadaísta literario en Zurich y luego se trasladó al París surrealista, pero conocía poco acerca de Loos. Un amigo me había comentado sobre el ensayo Crimen y Ornamento [2] y había visto algunas sillas expuestas en un museo de Viena.

Ahora descubría que vivo cerca de uno de los pocos edificios que construyó en París. Entonces, cuando me senté frente a la computadora, mirando por la ventana la esquina de la Avenida Junot y Simon-Dereure, pensé qué podría encontrar en la Web sobre Loos. ¿Dónde podría comenzar mi búsqueda? ¿Qué palabra clave podría usar para encontrar la información? Comencé por Google, en la actualidad el más efectivo buscador en Internet. Ahí me enfrenté, en primer lugar, con el problema de la ortografía. Los motores de búsqueda, como cualquier software, raramente saben qué deseamos decir, ellos sólo reconocen aquello que escribimos. Google se diferencia de su competencia al ofrecer el servicio “Quizás quiso decir” ante algún probable error de tipeo.

Pero en este caso, estaba buscando “Adolf” y “Adolph”, que son ambos teóricamente correctos. Empecé con “Adolph” y encontré una imagen de un retrato de Loos pintado por Oskar Kokoschka. Fue una falsa promesa, una pintura de un arquitecto pero ningún vínculo a su obra ni otra información sobre la pintura. Escribí “Adolf” y encontré mucho más.

Un ejemplo de la deficiencia de las “Authority Lists” [3] es el hecho de que un motor de búsqueda como Google sea incapaz de mostrar resultados de la búsqueda de una persona cuyo nombre puede ser escrito de diferentes maneras. Esta es una práctica común en bibliotecología, donde en un catálogo en que busquemos los trabajos de “Thomas Sterns Eliot” encontraremos T. S. Eliot, si buscamos “Samuel Clemens” hallaremos la obra de Mark Twain y para “Sami Rosenstock” encontraremos Tristan Tzara. Tipeando “Adolf Loos” encontré un sitio prometedor, ArtandCulture.com, que provee una útil colección de ensayos cortos y vínculos referentes a artistas, arquitectos, fotógrafos,

escritores y diseñadores, creación altruista de algunos mecenas del diseño en la Web durante el boom de las empresas punto com. Las influencias y relaciones de los artistas en una red hipertextual, facilitando encontrar un editorial sobre la vida y obra de Loos. Si hubiera empezado la búsqueda con “Oskar Kokoschka”, hubiera llegado a un artículo sobre el artista que retrató a Loos.

El editorial estaba seguido por una selección de recursos online que incluían la base de datos de Konstantin Melnikov creada por AGRAM [4], pero el vínculo a la página no funcionaba y en el navegador apareció el texto “Página no encontrada”. Tuve que apelar a mi experiencia profesional, editando el vínculo hasta conseguir la base de datos. La información estaba allí, pero no donde me llevaba el vínculo. ¿Por qué?

De hecho, ArtandCulture.com es una ciudad fantasma, un proyecto abandonado como muchos otros que desaparecieron. En el tiempo en el cual el editor del sitio hizo el vínculo y que yo hice click en él, la base de datos de AGRAM fue movida. No hay vínculo universal registrado para lidiar con este tipo de problemas, que en este caso era insignificante pero que puede tornarse muy serio al considerar la inmensa cantidad de artículos científicos que están disponibles en Internet.

Si estuviera investigando sobre procedimientos médicos y encontrara un vínculo a un artículo de trascendencia, me frustraría no poder acceder a la página porque no está disponible en esa dirección o porque el sitio fue absorbido por otra compañía. Procurando resolver este problema, se desarrolló un sistema [5] cuyo objetivo es “identificar e intercambiar propiedad intelectual en el entorno digital”. Este servicio permite registrar artículos a los cuales se les asigna un único número y así se los registra en la base de datos. De esta forma, el sistema sirve para mantener la “integridad” de los vínculos en el tiempo.

Volviendo a mi búsqueda, descubrí, en la base de datos de AGRAM, una fascinante colección de informaciones acerca de los edificios diseñados por Adolf Loos, creada por Rein Saariste en 1997 y que está en el servidor de la Universidad Tecnológica Delft, aunque ningún vínculo a ella pueda ser encontrado en el sitio de la Universidad. En AGRAM encontré una completa lista de todos los edificios diseñados por Loos, con las correspondientes fechas, ubicaciones, fotografías y planos. Refinando la búsqueda por año, encontré otros edificios en los cuales Loos trabajó en la época en la cual vivió. Clickeando en el edificio Rosenberg ¡me encontré leyendo sobre los planos de una construcción que estaría situada en el lugar que yo veía desde mi ventana!

Fue un momento especial e inusual para una experiencia de navegación por Internet. Observé en la pantalla de mi computadora la foto del modelo arquitectónico que Saariste [6] realizó sobre la base de los dibujos de Loos. Descubría que no puedo preguntarle a la Web “¿Qué puedes decirme acerca de la esquina de Junot y Simon-Dereure del 18° distrito de París?”. La Web no podría responderme “Te puedo mostrar una imagen de un edificio que fue diseñado 90 años atrás pero que nunca fue construido”.

Diseñando conexiones en la Matrix del Museo

Cuando tuve esta clase de experiencias, recordé el valor de las conexiones. El valor diferencial de la Web no está en encontrar cosas, sino en que los vínculos nos lleven a ellas, en una red hipertextual que podamos navegar. El entorno digital en el cual nos movemos, hace que no pongamos atención en las conexiones más obvias.

La trascendencia del problema no debe ser buscada en los pequeños sitios y páginas individuales, sino en los sitios de las grandes corporaciones donde la ausencia de arquitecturas de la información cuidadosamente planificadas ocasiona que ciertos contenidos permanezcan escondidos en los espacios de información que albergan los sitios.

Un ejemplo de un muy eficiente uso del “Enterprise Web Space” [7] es el sitio del Museo Metropolitano de Arte de Nueva York [8] (Met), donde encontramos excelencia en el diseño, la arquitectura de la información, la usabilidad y el desarrollo de contenidos. Utilizo “enterprise” para referirme a cualquier gran y compleja institución. En economía y negocios remite a una gran compañía o corporación con múltiples locaciones y divisiones. El problema que se afronta es el mismo cuando hablamos de un importante museo con muchos departamentos o de una corporación multinacional con diversas líneas de productos y con presencia en mercados regionales.

El “Enterprise Web Space” consiste en todos los elementos que facilitan el conocimiento, por parte de los usuarios, de una organización pública o privada en Internet. Los elementos incluyen: sitios, aplicaciones, contenidos, newsletters y todo material o servicio al que se pueda acceder desde la Web. Cada componente del “Enterprise Web Space” es conducido por diferentes grupos, con distintos presupuestos. Pero considerar este espacio como un grupo, un sitio o un grupo de sitios, constituye en gran error. El “Enterprise Web Space” es parte de una amplia estructura integrada definida en el marco de una perspectiva del usuario quien al visitar el sitio de un museo, un espacio de información único, puede encontrar información acerca de una muestra, una publicación, su pintura favorita, un recurso educacional, una referencia histórica, combinadas de cualquier forma.

Si bien el sitio del Met tiene una sola página de inicio, está construido por muchos sitios, bases de datos, servicios de comercio electrónico y colecciones de imágenes, cuyos contenidos son administrados por diferentes objetivos, con fines que difieren entre sí. Para hacer este “espacio” útil, debemos diseñar una adecuada Arquitectura de la Información [9] que requiere crear estructuras centradas en el usuario con las respectivas relaciones que mantendrán.

La home page del Met presenta diversas formas de organización paralelas. En la barra izquierda tenemos una lista jerárquica dividida en secciones: Obras de arte, tienda virtual, miembros, información para visitantes, Mi museo Met (un servicio personalizado) y noticias. La parte central superior muestra las tres exhibiciones actuales y en la parte central inferior se repiten los ítem que aparecen en la barra de navegación izquierda sumando un ítem dedicado al museo para los niños y otro para recursos educativos de arte.

Si pensamos al Met como una colección de arte, es lógico empezar con “Colecciones”, donde nos encontramos con los departamentos que integran el museo, información crítica para la institución pero de menor relevancia para quien visita el sitio. Si seleccionamos “Features”, el arte del mundo es unificado una vez más y podemos acceder a trabajos, adquisiciones recientes, recursos educativos y tour virtual.

Una vez que seleccionamos un Departamento, por ejemplo el de Pintura y Escultura

Americana, accedemos a colecciones destacadas, bases de datos, ensayos, mapa de las galerías, listado de publicaciones, recursos y exhibiciones donde podemos encontrar vínculos a tópicos especiales, muestras en curso, sitios especialmente realizados para determinados eventos e información sobre exhibiciones que ya finalizaron pero que siguen online.

La actual muestra sobre el pintor americano Gilbert Stuart tiene su propio sitio, que presenta imágenes de la exhibición organizadas por sala temática. Podemos seleccionar una obra en particular, por ejemplo un retrato de George Washington pintado por Vaughan, y leer una pequeña descripción. Pero no existen vínculos a las obras fuera del marco del sitio de la exhibición sin que medien razones técnicas, salvo una cuestión de diseño.

El Vaughan que seleccionamos es un préstamo de la Galería Nacional en Washington [10] y podemos encontrarlo en el sitio de ese museo con una presentación diferente. El sitio es otro enorme “Enterprise Web Space” y es conducido por diferentes grupos, con sus fortalezas y debilidades. Entre las primeras está la integración entre la obra del artista en el marco de una importante base de datos que contiene el sitio. La pintura que seleccionamos está vinculada a todos los datos relevantes de la misma, como su descripción, bibliografía, historial de exhibiciones, comentarios y pinturas relacionadas, entre otras. Lo más destacado es que el nombre del artista está vinculado a toda su obra, en forma útil y simple, seguramente a través de una “Authority List”.

Retornando al sitio del Met, en Works of Art > Timeline of Art History tenemos un enorme sitio organizado sobre la base de un criterio temporal (20.000 a.C. hasta el presente) y otro geográfico. En este sitio las conexiones son fundamentales para el contenido volcado, ya que apunta a una audiencia general. La página sobre la exhibición de Gilbert Stuart presenta una selección de pinturas y ensayos con vínculos a otros artistas y temáticas. La barra de navegación izquierda del “Timeline...” presenta un listado completo de temas, mapas y recursos. En la barra de navegación derecha hay conexiones a materiales educativos en su sitio “Explore & Learn” como también vínculos a exhibiciones especiales, incluyendo la que visitamos.

Merece destacarse que si bien desde la página “Timeline...” estamos conectados no solamente a todo el sitio sino a otros relacionados, desde la página sobre la exhibición de Gilbert Stuart (En la sección Colecciones) estamos aislados de cualquier información provista en el “Enterprise Web Space”.

La búsqueda que encuentra cosas en todos lados

Diseñar un sistema de vínculos, en el marco de un sistema de navegación y etiquetas, es sólo una parte de la solución para todo sitio corporativo de Internet. Otra cuestión crítica es la arquitectura de la información diseñada para el proceso de búsqueda, que incluye cómo los resultados son organizados y presentados. La arquitectura de la información no está limitada por los avances tecnológicos, en los motores de búsqueda. La búsqueda en “texto completo” [11] es la más utilizada en los grandes sitios.

La mayoría de los usuarios de Internet no son especialistas en computación y, por ello, no les resulta interesante el “detrás de la escena” de los motores de búsqueda. Un “Enterprise Web Space” suele tener diversos índices de búsqueda, porque el contenido desestructurado de las páginas del sitio difieren de la base de datos de la colección del museo, y porque el contenido de la colección de imágenes digitales es distinto de los

contenidos y servicios del sistema de compra en línea.

Pensar en desarrollar diferentes opciones de búsqueda para cada contenido es una solución insatisfactoria. Primero, porque significa asumir que el usuario tipo no encontró lo que buscaba porque hizo la pregunta correcta en el lugar equivocado. Segundo, presume que el usuario sabrá encontrar el lugar correcto. También es una mala solución listar todos los resultados en una lista única. Hay diferencia entre libros en venta, una serie de pinturas y materiales docentes, entonces ¿por qué descartar el contexto?

El sitio del Met ofrece una buena solución, al consultarle al usuario dónde quiere realizar la búsqueda y donde cada opción ofrece un útil conjunto de respuestas, integrando los contenidos dinámicos y los estáticos del sitio. Nos resultará natural encontrar una clase si elegimos “Explore & Learn”, o realizar una compra si optamos por la tienda del museo. El buscador del calendario presenta una completa lista de eventos, con un útil sistema de etiquetas que nos permite distinguir si se trata de una conferencia, una actividad académica, o una charla en la galería.

Podemos volver al sitio de la Galería Nacional para encontrar otro ejemplo. Si quiero encontrar “Gilbert Stuart” estoy forzado a elegir entre buscar en el sitio o solamente dentro de “Colecciones”, y si deseamos buscar en ambos debemos tipear la palabra clave dos veces. Debemos entrar a la sección “Recursos”, para acceder a contenidos que estén en la base de datos de la biblioteca que requiere búsquedas separadas para los ítem clasificados como libros, diapositivas, fotos, entre otros. Si bien el catálogo de la biblioteca difiere en cuanto a contenido del resto de las páginas, forman parte del mismo espacio. Entonces, ¿por qué tienen que existir partes de contenido “escondidas”?

Un muy buen ejemplo de una arquitectura de la información para los motores de búsqueda es el del sitio de la biblioteca nacional australiana (NLA) [12] que, como la mayoría de otros similares, tiene más de un catálogo. Diferentes bases de datos contienen información sobre manuscritos, fotografías, grabaciones, pinturas y libros. También contiene colecciones digitales cuyas obras datan desde el siglo XVIII hasta la actualidad, las cuales podían ser encontradas en sitios específicos cada uno con su propio motor de búsqueda.

En la actualidad, la búsqueda se unifica en una interfaz que presenta los ítems encontrados en las diferentes secciones, donde los primeros resultados de cada categoría incluyen pequeñas imágenes de las colecciones digitales y la posibilidad de optar por ver los resultados más relevantes o el listado completo. De una forma rápida, nos es presentada la información que podemos encontrar y se nos recuerda la variedad de recursos que disponemos.

Esta solución no requiere de una gran inversión tecnológica sino de lograr un buen producto desde la imaginación del arquitecto de la información y sus conocimientos aplicados en diseño centrado en el usuario. Las interconexiones realzan el valor de la experiencia del usuario, sin dejar de lado la importancia de mantener bases de datos separadas. Trabajando en esta dirección, las organizaciones tendrán beneficios en la relación con sus usuarios, maximizando el poder de Internet como medio de comunicación.

NOTAS

* Este artículo fue publicado originalmente en el número 14 de la revista polaca 2+3D, en enero de 2005, por lo cual algunos ejemplos o citas pueden no encontrarse online al momento de la edición de este dossier.

[1] El artículo está disponible online (inglés) en

<http://www.theatlantic.com/doc/194507/bush>

[2] El ensayo fue publicado en 1907 en la revista Das Andere.

[3] Las "Authority Lists" pertenecen al grupo de los vocabularios controlados en Internet y refieren a una lista de términos preferidos y/o valores aceptados, que no incluyen variantes o sinónimos.

[4] www.agram.fr.

[5] El "Digital Object Identifier System" fue desarrollado por la Fundación Internacional DOI. <http://www.doi.org/>

[6] R. Saarite documentó el trabajo de arquitectos que influyeron el desarrollo de la arquitectura en el uso del espacio, incluyendo a Le Corbusier, Loos, Jujol y otros.

<http://agram.saarite.nl>

[7] <http://www.kahnplus.com/ourwork/en/space.htm>

[8] <http://www.metmuseum.org>

[9] Paul Kahn expuso sobre el tema en ICHIM 2004, desarrollado en Berlín.

<http://www.ichim.org/jahia/Jahia/cache/offonce/pid/14>

[10] <http://www.nga.gov>

[11] Técnica utilizada en todos los documentos almacenados en un sitio, basada en las características del texto completo de los mismos.

http://en.wikipedia.org/wiki/Full_text_search

[12] <http://www.nla.gov.au/>